



# **NAVAL POSTGRADUATE SCHOOL**

**MONTEREY, CALIFORNIA**

## **THESIS**

**AN ANALYSIS OF MODELING SUCCESS IN EXPLOSIVE  
ORDNANCE DISPOSAL TRAINING**

by

Trevor J. Ritland

March 2010

Thesis Advisor:  
Second Reader:

Samuel E Buttrey  
Robert A. Koyak

**Approved for public release; distribution is unlimited**

<b>REPORT DOCUMENTATION PAGE</b>			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> March 2010	<b>3. REPORT TYPE AND DATES COVERED</b> Master's Thesis	
<b>4. TITLE AND SUBTITLE</b> An Analysis of Modeling Success in Explosive Ordnance Disposal Training			<b>5. FUNDING NUMBERS</b>	
<b>6. AUTHOR(S)</b> Trevor J. Ritland				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Naval Postgraduate School Monterey, CA 93943-5000			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Center for Explosive Ordnance and Diving Panama City, Florida			<b>10. SPONSORING/MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b> The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. . IRB Protocol number _____.				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited			<b>12b. DISTRIBUTION CODE</b>	
<b>13. ABSTRACT</b>  This thesis is a follow-on study to the Master of Business Administration (MBA) project, <i>Grade Point Average as a Predictor of Success at Explosive Ordnance Disposal Training</i> , completed in December 2009 by Lieutenant Sarah Turse and this author. The purpose of this thesis is to analyze, develop and provide a more accurate student graduation prediction model than the current model in place at Naval School Explosive Ordnance Disposal (NAVSCOLEOD). The school's current model was produced five year ago using ordinary linear regression. This outdated model was compared to the new model generated in this study using statistical techniques such as receiver operating characteristic (ROC) curves and the Hosmer-Lemeshow test. Our analysis finds that the student's branch of service, GPA, and the division in which the student failed each significantly impact predicting a student's future. We also find that the interaction between GPA and division also significantly impacts the prediction. Finally, we conclude that using a logistic regression instead a linear regression captures the binary output (graduated or did not graduate) better. Our improved model increases the prediction probability by roughly 2 percent using student data from 2004 to 2008.				
<b>14. SUBJECT TERMS</b> Master of Business Administration (MBA), Grade Point Average (GPA), Naval School Explosive Ordnance Disposal (NAVSCOLEOD), Receiver Operating Characteristic (ROC), Hosmer-Lemeshow test			<b>15. NUMBER OF PAGES</b> 75	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b> UU	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)  
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release; distribution is unlimited**

**AN ANALYSIS OF MODELING SUCCESS IN EXPLOSIVE ORDNANCE  
DISPOSAL TRAINING**

Trevor J. Ritland  
Lieutenant, United States Navy  
B.S., United States Naval Academy, 2002

Submitted in partial fulfillment of the  
requirements for the degree of

**MASTER OF SCIENCE IN OPERATIONS RESEARCH**

from the

**NAVAL POSTGRADUATE SCHOOL  
March 2010**

Author: Trevor J. Ritland

Approved by: Samuel E. Buttrey  
Thesis Advisor

Robert A. Koyak  
Second Reader

Robert F. Dell  
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

## ABSTRACT

This thesis is a follow-on study to the Master of Business Administration (MBA) project, *Grade Point Average as a Predictor of Success at Explosive Ordnance Disposal Training*, completed in December 2009 by Lieutenant Sarah Turse and this author. The purpose of this thesis is to analyze, develop and provide a more accurate student graduation prediction model than the current model in place at Naval School Explosive Ordnance Disposal (NAVSCOLEOD). The school's current model was produced five year ago using ordinary linear regression. This outdated model was compared to the new model generated in this study using statistical techniques such as receiver operating characteristic (ROC) curves and the Hosmer-Lemeshow test. Our analysis finds that the student's branch of service, GPA, and the division in which the student failed each significantly impact predicting a student's future. We also find that the interaction between GPA and division also significantly impacts the prediction. Finally, we conclude that using a logistic regression instead a linear regression captures the binary output (graduated or did not graduate) better. Our improved model increases the prediction probability by roughly 2 percent using student data from 2004 to 2008.

THIS PAGE INTENTIONALLY LEFT BLANK

## TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	OBJECTIVES .....	1
B.	BACKGROUND AND MOTIVATION .....	1
II.	LITERATURE REVIEW .....	11
A.	NAVSCOLEOD INSTRUCTION 5420.1U .....	17
III.	METHODOLOGY .....	27
A.	MODEL REGRESSION .....	27
B.	HISTORICAL DATA.....	27
C.	MODEL PREDICTORS .....	28
D.	ADDITIONAL MODEL PREDICTORS .....	28
	1. Branch of Service .....	29
	2. Interaction Among the Predictors .....	29
E.	COMPARISON OF MODELS.....	29
	1. ROC Curves .....	29
	2. Hosmer_Lemeshow Test .....	31
	3. Model Subset Comparison .....	33
IV.	DATA ANALYSIS .....	35
V.	RESULTS .....	39
VI.	CONCLUSION .....	51
A.	RECOMMENDATIONS.....	52
	LIST OF REFERENCES .....	53
	INITIAL DISTRIBUTION LIST .....	57



THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF FIGURES

Figure 1.	Backward-Looking Analysis .....	23
Figure 2.	Forward-Looking Analysis .....	24
Figure 3.	Prediction vs. Outcome Table.....	30
Figure 4.	ROC Curve Analysis (From Lahanas, 2010).....	31
Figure 5.	Logistic Function Graph .....	36
Figure 6.	2004–2008 Divisional Setbacks.....	37
Figure 7.	ROC Curve of New and Current Models.....	46
Figure 8.	ROC Curve of New (Without Branch) and Current Models .....	46
Figure 9.	Hosmer-Lemeshow Plot of the New Model .....	47
Figure 10.	Hosmer-Lemeshow Plot of the New Model (Without Branch).....	48

THIS PAGE INTENTIONALLY LEFT BLANK

## LIST OF TABLES

Table 1.	Allowable Model Accuracy Tolerance Based on Sample Size .....	18
Table 2.	ARB Model Probability Matrix .....	20
Table 3.	FY05 ARB Model Validation.....	20
Table 4.	FY05 ARB Model Validation With an Adjusted Decision Threshold .....	21
Table 5.	2004–2005 Model Prediction vs. Student Outcome (Threshold = .5) .....	22
Table 6.	2004–2005 Backward-Looking Analysis .....	22
Table 7.	2004–2005 Forward-Looking Analysis .....	24
Table 8.	Model Score Subsets.....	32
Table 9.	Rank-Ordered, Graduation Rate Comparison Example.....	33
Table 10.	New Model Equation .....	41
Table 11.	New Model Equation Without Branch .....	42
Table 12.	2004–2008 Current Model Prediction Outcomes .....	43
Table 13.	2004–2008 Current Model Forward-Looking Analysis .....	43
Table 14.	2004–2008 New Model Prediction Outcomes.....	44
Table 15.	2004–2008 New Model Forward-Looking Analysis .....	44
Table 16.	2004–2008 New Model (Without Branch) Prediction Outcomes.....	45
Table 17.	2004–2008 New Model (Without Branch) Forward-Looking Analysis.....	45
Table 18.	Two-Way Table Overall Prediction Probability .....	45
Table 19.	Hosmer-Lemeshow Test Comparison.....	48
Table 20.	Rank-Ordered Graduation Rate Comparison.....	49

THIS PAGE INTENTIONALLY LEFT BLANK

## **EXECUTIVE SUMMARY**

Since the beginning of the wars in Iraq and Afghanistan, there has been a rapid increase in demand for military Explosive Ordnance Disposal (EOD) technicians. The military's EOD units have been the main assets in countering the enemy's number one weapon of choice, the improvised explosive device (IED). As a cheap and highly effective weapon, the IED is responsible for thousands of military and civilian casualties along with creating an unstable and unsecure environment. As a result, coalition forces have expanded their counter-IED efforts by increasing the number of EOD units and technicians in the field. The United States is no exception and has been the main provider of EOD assets.

To support this demand for additional manning, the U.S. has increased the number of students it sends to EOD training. While this seems like a simple solution, there are a limited amount of resources the military has to train its EOD assets. Over the past seven years, there has been a student increase of 44 percent. As the single source provider, the Navy manages the one basic EOD School that supports all four services—the Army, Air Force, Navy and the Marines. EOD School is an intensive, 42-week training program that educates students on types of explosives and their various types of employments. EOD School is challenging and generally attrites roughly 25–30 percent. As a result of the increase in students, the number of classes each year has grown, the class sizes have grown, the staff to support the students has grown, and the amount of training equipment has grown. While most areas of EOD School have been able to keep up with the demand, there are certain parts that have trouble maintaining the current tempo.

When a student fails the same exam twice while at EOD School, the student is removed from training and sent to an Academic Review Board, or ARB. This is called an academic setback. The ARB is a process to decide whether a student should be kept in training and “rolled back” to a new EOD class, or removed from training completely. There are roughly six members to an ARB. This process is intensive and time

consuming. Not including the counseling and paperwork, boards generally last a half an hour for each student. Before the wars in Iraq and Afghanistan began, these boards convened for roughly 40 percent of EOD students. While the percentage has remained the same, the number of ARBs has drastically increased. This has caused a bottlenecking effect of students waiting for a board. Student delays have lasted up to weeks. Furthermore, the success of the board decision has impacted the efficiency of EOD School. There are some students who “roll back” two or three times before being removed from school.

As a solution to the ARB problem, EOD School developed a decision tool, based on historical data, that uses a student’s information to predict, at the point of the student’s first academic setback, whether he or she will graduate. This model allows EOD School to forgo a full ARB and increases the throughput of students who have a setback. Less time is wasted for the board members, and the students awaiting those boards. While this was a good model five years ago when it was first developed, changes in the school’s curriculum are cause for concern whether the model still is accurate.

To find out, Turse and Ritland (2009) used the most recent student information to verify whether there was any deviation from the accuracy the School’s instruction claimed. After completing their research, they found a slight degradation in the model’s performance and an error in the instruction that provides misleading information on the overall accuracy of the model. This misleading information gives the perception that the model has a 95 percent prediction success rate. When the same data was applied to the more appropriate perception, the success rate dropped to roughly 70 percent.

Furthering Turse and Ritland’s (2009) research, this thesis analyzed the recent data to create an improved prediction tool to aid the decision-making process of the ARB.

## **LIST OF ACRONYMS AND ABBREVIATIONS**

ARB	Academic Review Board
ASVAB	Armed Service Vocational Aptitude Battery
CISO	Chief Information Security Officer
CO	Commanding Officer
CPI	California Psychological Inventory
EOD	Explosive Ordnance Disposal
FPR	False Positive Rate
GLM	Generalized Linear Model
GPA	Grade Point Average
HPI	Hogan Personality Inventory
JIEDDO	Joint Improvised Explosive Device Defeat Organization
LOGIT	Logit Function
LSI	Learning Styles Inventory
LT	Lieutenant
MBA	Master of Business Administration
MIDAS	Multiple Intelligences Development Assessment Scales
NAVEDTRA	Naval Education Training
NAVSCOLEOD	Naval School Explosive Ordnance Disposal
NAVSCOLEODINST	Naval School Explosive Ordnance Disposal Instruction
RFF	Request for Forces
ROC	Receiver Operating Characteristic
SDS	Self Directed Search
SME	Subject Matter Expert
TO	Training Officer
TPR	True Positive Rate
USN	United States Navy
XO	Executive Officer



THIS PAGE INTENTIONALLY LEFT BLANK

## **I. INTRODUCTION**

### **A. OBJECTIVES**

This study continues the analysis done by Turse and Ritland (2009) on the academic prediction model used at the U.S. military's Explosive Ordnance Disposal (EOD) School. Although that project was limited to validating the academic prediction model and discussing its limitations, this study will use other predictors to maximize the school's ability to determine a student's success probability. Using the same historical data provided by EOD School, this study included breaking the data down into the individual components such as service, age, rank, and where the student is in his or her training, and looks at the interaction the individual components have with one another. This study seeks to benefit the school in several ways:

- Provide insight to aid in making important academic decisions;
- Minimize the amount of time students spend waiting to be employed;
- Minimize the number of hours instructors spend conducting student review boards and associated tasks;
- Minimize the strain on the school in supporting its students;
- Save EOD School and the Navy money.

### **B. BACKGROUND AND MOTIVATION**

Over the past several years, the wars in Iraq and Afghanistan have caused the U.S. military to change and adapt to an evolving asymmetric enemy. In a military dominated by a conventional military mindset, the U.S. has had to quickly shift its focus to counter the prevalent unconventional threat. To do this, thousands of additional specialized troops have been required to halt the enemy's progression in combating coalition forces. In 2007, President Bush deployed an additional 21,500 troops to Baghdad and Al-Anbar province, which increased total U.S. forces in the country to 135,000 (Bender, 2007). In

2009, President Barack Obama ordered deployment of an additional 17,000 troops to Afghanistan in February (Barnes & Miller, 2009) and 30,000 more nine months later (Schmitt, 2009). The increase of forces impacted not only the military units sent directly into combat, but also the training commands required to support the additional forces. Because the enemy has thrived on unconventional warfare techniques, like the use of suicide bombers and Improvised Explosive Devices (IEDs), EOD technicians continue to be in high demand (Atkinson, 2007). These new force requirements have strained the EOD community military wide and the training units that support them.

Even today, the EOD communities in each of the military services are undermanned. Currently, the enlisted manning for Navy EOD is at 95 percent of its authorized strength. Though 95 percent may seem reasonable, Navy EOD's Zone A manning (sailors who have 17 months to six years of service) is only 75 percent. The Air Force, Army and Marine Corps are respectively manned at 80 percent, 84 percent and 78 percent, according to personal correspondence with their service representatives. As IEDs continue to be the number one killer in the wars in Iraq and Afghanistan, keeping the military's EOD assets at capacity continues to be a challenge (JIEDDO, 2008). It is important to increase and maintain EOD strength as EOD technicians provide the main countering force against terrorists' number-one weapon of choice worldwide, the IED (Atkinson, 2007). IED construction is simple, requires limited skill, and gives the terrorist "the ability to conduct spectacular attacks for relatively small investment. IEDs continue to provide the enemy with inexpensive, stand-off, precision weapon systems that often provide the attacker with near total anonymity" (JIEDDO, 2008).

Overall, the numbers of EOD responses and deaths due to IEDs have steadily increased from when the wars in Iraq and Afghanistan first started. In Iraq, U.S. deaths due to IED attacks increased from 31 percent of the total fatalities in 2004 to 66 percent of the total fatalities in 2007. By September of 2007, a total of 1509 Americans had died in Iraq due to makeshift roadside bombs (Hagey, 2007). EOD units are the primary assets to respond to IED threats. In 2006 alone, the total number of Joint EOD responses in combat zones totaled 20,890 (Wehmeyber, 2007). By the end of 2007, Iraq began to see a slight decline in monthly attacks. In 2008, IED incidents accounted for only 40

percent of attacks on coalition forces in Iraq, reaching their lowest levels since 2003. “The total number of IED attacks in September 2008 was 33 percent of September 2007 and 22 percent of September 2006 levels” (JIEDDO, 2008). While Iraq showed signs of improvement, Afghanistan however, rapidly got worse. In 2008, 3,376 IEDs detonated or were detected before blowing up, a 45 percent increase from 2007. One year later, IED attacks increased a further 146 percent (Brook, 2009). As the primary asset to respond to an IED threat, EOD forces have been strained with multiple deployments and extensions to satisfy the needs of the theater Component Commander (COCOM) to respond to these threats and help protect his units.

EOD technicians support the battlefield commander by addressing the primary threats to the coalition forces. Their missions include direct action assault support, IED response, unexploded ordnance response, route clearance, subject matter expert (SME) advice, and post-blast analysis. The range of the EOD mission set continues to expand as adversaries adapt and incorporate both low-tech and more sophisticated technologies to wage war, e.g., bombs made from fertilizers, internet recruiting, and cell-phone bomb activation. As a result, deployment rates for military EOD are at an all-time high, with some services reaching a dwell time of 1:1. In other words, EOD technicians are spending as many days deployed as they are at home (Wehmeyer, 2007). Despite the various incentives such as early promotion, re-enlistment bonuses and special duty assignment pays to EOD technicians, community manning remains insufficient to meet requirements. While each service has increased the number of students at EOD School to counter these manning shortfalls, the high attrition rates have contributed to persistent low manning strengths.

The process to become an Explosive Ordnance Disposal technician is a long and arduous one. Turse and Ritland (2009) described this process in detail.

Each EOD candidate, regardless of service affiliation, must complete an intensive training curriculum at Naval School Explosive Ordnance Disposal (NAVSCOLEOD) located on Eglin Air Force Base Florida. Challenging in both physical and mental demands, the program lasts at least 42 weeks and possibly longer, depending on military service and training setbacks. The Navy extends its training over 68 weeks to include dive school, underwater ordnance, airborne school, and tactical training.

This is due to the differences in mission requirements (Navy EOD, 2009). An academic setback occurs when a student cannot complete the required learning objectives for a specific area of study and must repeat the training. The attrition rates vary by service with the Air Force typically having the highest and the Marine Corps the lowest. Overall the average attrition rate is around 27 percent over the last five years.

EOD School itself consists of 12 divisions: 1) Core I, 2) Demolition, 3) Reconnaissance, 4) Tools and Methods, 5) Core II, 6) Ground Ordnance, 7) Air Ordnance, 8) Improvised Explosive Devices, 9) Biological and Chemical Weapons, 10) Nuclear Weapons, 11) Weapons of Mass Destruction and 12) Underwater Ordnance. Each of these divisions varies in length and complexity. Typically Ground Ordnance and Air Ordnance divisions are the more challenging divisions. Underwater Ordnance division is for naval personnel only. The student must be dedicated and focused to successfully pass each division and graduate. The school's curriculum challenges its students in order prepare them for often intense stresses as an EOD technician. As stated on the Navy's EOD website, 'Upon graduation, EOD technicians are equipped with the skills necessary to render safe and dispose of high explosive material in permissive and non-permissive environments.'

The increasing demand for EOD personnel in Iraq and Afghanistan, along with the existing manning shortfalls, has caused a strain on the military EOD community. One solution is to substantially increase student throughput at NAVSCOLEOD. In just four years (2004 to 2008), billets at EOD School have increased from 777 students to 1122 students, a 44 percent increase (Andrea, 2009). These additional students will benefit the EOD community by assisting with the mission requirements around the world and increasing the manning for deployment relief.

To accommodate additional students, EOD School has been required to make a few changes. These changes include increasing the number of EOD classes each year, increasing the number of instructors, and increasing sizes of the classes. This expansion has forced the school to acquire more training equipment and facilities.

Another challenge the school faced was the increase in the number of students facing an academic setback. When a student experiences an academic setback, he or she awaits an Academic Review Board (ARB). The ARB is designed to evaluate the student's academic progress and make recommendations concerning student training

potential. A setback is administered when students do not meet training objectives, most commonly evidenced in a written or practical exercise test failure. An ARB may be convened at any time if the division officer feels the student has become so far behind that training objectives will not be met for the division, or the student appears to reflect a safety hazard (NAVSCOLEOD, 2008).

The number of setbacks throughout school is substantial. Over the 42-week curriculum, approximately 40 percent of students receive at least one setback in training. In years 2002–2007, there were 1391 out of 3597 students (roughly 39 percent of students) who experienced at least one academic setback (Andrea, 2009). While going through school, a student will get a second opportunity on an exam if he or she fails an exam. After two failures on the same exam, the student is removed from training while the instructors and staff in an ARB evaluate whether the student will be able to complete the assigned training. The ARB determines if the student will be allowed to repeat the division, or will be removed from training permanently. Typically, a student who experiences one academic setback is allowed to return to training in a later class. However, if a student requires a second ARB for academic deficiency, the likelihood of graduating is low. Extensions in training are financially costly as they increase a student's total time in training. These extensions are not guarantees for students to complete training. If they fail to graduate, it is even more costly as there is no return on investment on capital or manpower.

The ARB board serves many purposes. These include helping students solve problems that may prevent successful completion of training, as well as identifying which students are capable of completing the training. The ARB also identifies students who are unwilling or unable to complete EOD training. Based on the findings, the board makes decisions and recommendations concerning the future of each of the academic setback students (NAVSCOLEOD, 2008). The school's policy states:

The possible ARB Recommendations include:

1. Continue with Class: Continuation of training in the present class with or without remediation.

A. Without Remediation: The student is not required to take a retest and the student has met all training objectives.

B. With Remediation: The student is required to take a retest. The student has not met training objectives and must successfully pass retest prior to completion of division.

2. Setback: The student receives an extension of training with remediation. The student will repeat all or part of the current division or previous division curriculum as recommended by the ARB. Students will normally be set back only to repeat the training objectives that have not been satisfactorily demonstrated. If repeating additional training objectives that precede the failed training objective is a remediation method that will benefit the student, the ARB may recommend it.

3. Drop from Training: A student has not met training objectives and should be permanently removed from training. When the ARB recommends a drop from training, the student must have demonstrated unwillingness or inability to continue the training.

The school's instruction lists the following people who have authority and responsibility in the academic setback process:

1. Testing Officer. Ensure that the student is tested in accordance with instruction.

2. Instructor. Ensure NAVSCOLEOD Form 1610/1 Sections I through VII are completed properly. Once all information is correctly filled out, the package will be forwarded to the Division Officer for action.

3. Division Officer. Ensure the student is counseled on his or her failure; fills out NAVSCOLEOD Form 1610/1, check previous sections for correctness and forward to the training officer for action.

4. Training Officer. Conduct student interview and review student's academic history. For a student's initial setback, the Training Officer (TO) will determine if the student meets the requirement for an academic setback or drop from training per instruction. For all other setbacks, an ARB will be convened. The Training Officer will ensure ARB packages are complete, deviation notices updated, and students are present prior to ARB convening in proper uniform. The Training Officer will work with CISO to ensure board composition is in compliance with instruction, based on service component.

5. ARB Board Chairman. Ensure boards are conducted per instruction. The student record with disposition recommendation will be forwarded to the Commanding Officer (CO) or Executive Officer (XO) by the Training Officer and via the Service Detachment Commander. The Executive Officer will retain final decision for ARB recommendations, which are agreed to by the Service Detachment Commander. All others will be forwarded to the Commanding Officer for final decision. In these cases, the Service Detachment Commander or his designated representative may present the board package to the CO. The Training Officer will take necessary action to effect the student's disposition. If the student is dropped from training, he or she shall be turned over to the Training Support Officer or Service Detachment Commander. The second and all subsequent ARBs will consist of a chairman and at least two additional service members who shall be certified instructors.

6. Detachment Commanders. Service Detachment Commanders and/or Liaison Officers will be notified by the Training Department of all impending boards. Commanders will review the ARB package and make recommendations to the Commanding Officer.

7. Commanding Officer. Review all completed ARB packages and exercise final disposition authority in all cases unless the Service Commander disagrees with the ARBs recommendation. (NAVSCOLEOD, 2008)

The school's instruction, NAVSCHOLEODINST 5420.1U, also states:

Setbacks are categorized as Academic or Non-Academic, depending on the circumstances. Non-academic setbacks may occur when the student is unable to complete training due to illness or special circumstances outside the control of the course or the student. Academic setbacks occur after the failure of the first retest. The Training Officer, as a result of unsuccessful remediation and retesting, may grant initial academic setbacks. Remediation efforts may include supplemental examinations by the Division Officer with approval of the Training Officer. Supplemental examination will only be given if an administered test is deemed invalid due to technical information or instructor error. The Training Officer will inform CISO when a supplemental test is directed. CISO will take appropriate action per NAVSCOLEOD instruction. If remediation can be achieved in any way other than setback, it shall be considered first. Students will be set back only when the training objectives have not been satisfactorily demonstrated (NAVSCOLEOD, 2008).



The academic setback process is a lengthy, time-consuming, and paperwork-intensive procedure. While divisional instructors spend hours counseling each student who receives an academic setback, the majority of the time is spent completing an academic review board. Academic review boards are made up of six people: a chairman, four service representatives, and the student who failed an exam twice. If the same student needs a second academic board, two additional board members will be included. Each of these boards requires at least 30 minutes of review and decision-making, and is composed of individuals who are pulled from their primary jobs to sit as board members.

As the school increases the number of students to fill manpower shortages in the EOD community, the number of ARBs has substantially increased. Consequently, this increase has caused a bottlenecking of students awaiting an ARB. To mitigate this problem, in 2006, the school created a prediction model based on historical data to expedite the ARB process. This model used the student's grade point average (GPA) at the point of academic setback, along with the division in which the student failed, to predict whether or not he or she would succeed. With an acceptable level of error, the school found the model to be very effective. For example, in the months from October 2008 until March 2009, 306 students were subjected to an ARB. Had full academic review boards been held for each of these 306 students, as was done before the new Training Officer Setback policy (see section II), at least 918 man-hours would have been required to handle this process (306 boards x 6 people/board x 30 min/board). This number further breaks down into 765 hours lost to instructors and 153 hours lost to students. Instructors, therefore, lost over 95 days to review boards in the last six months ( $765 \text{ hours} / 8\text{-hours/day} = 95.64 \text{ days}$ ) and students lost 19 days to review boards in the last six months ( $153 \text{ hours} / 8 \text{ hours/day} = 19.13 \text{ days}$ ) (Turse & Ritland, 2009).

Turse and Ritland (2009) estimated that the dollar cost associated with workdays missed in order to attend ARBs was around \$26,000 for six months. This amount does not include the time and costs for counseling, paperwork and remediation. Additionally, time and costs are lost due to setback students waiting to be "recycled" back into training with the next available class. Unfortunately, as the student population at the schoolhouse

increases, time and money lost will continue to accumulate; however, the costs can be minimized by utilizing “tools” such as the prediction model.

Turse and Ritland’s (2009) study analyzed this graduation prediction model. At the time, the model was four years old and needed to be updated. Their results found the model to be slightly out of standards. They also concluded that the school was using the wrong measure of accuracy of the model. To fix this, they provided a new perspective in interpreting the data in which they felt more clearly portrays the accuracy of the model. This is discussed further in Chapter II.

THIS PAGE INTENTIONALLY LEFT BLANK

## II. LITERATURE REVIEW

The study by Turse and Ritland (2009) reviewed the literature on training students to be EOD technicians. The first study Turse and Ritland looked at was called *Commonalities in an Uncommon Profession: Bomb Disposal*, conducted by Bundy from the Technical Support Working Group, Explosives Detection Subgroup. In this study, Bundy stressed the importance of having highly trained, responsive and consistent military or civilian bomb disposal technicians in any environment. The intent of his study was to identify, through a series of analytical tests, if intelligence strengths and learning styles are common among the professionals in the EOD organization. He focused on identifying the cognitive characteristics that most successful EOD technicians possess. More specifically, he wanted to “investigate the extent to which individual learner preferences, as measured by learning styles and multiple intelligences, impact the effectiveness of bomb disposal training” (Bundy & Sims, 2007).

Bundy used the Canfield Learning Styles Inventory (LSI) and Multiple Intelligences Developmental Assessment Scales (MIDAS) to analyze learner preferences. Using these techniques, he determined that EOD technicians tend to share common traits that therefore could be used as predictors for academic success. Although the research identified these traits, it was limited and did not address factors such as the psychological and neuropsychological characteristics of EOD technicians.

Bundy conducted his research by sampling a wide demographic of EOD technicians, both civilian and military to determine his results. He concluded that commonalities existed despite race, age, gender, education, and military affiliation. Bundy noticed that bomb technicians tend to prefer specific learning styles while showing aversion to others. Furthermore, he noticed that EOD technicians shared similar strengths and weaknesses in certain types of intelligence. Uncovering this information, Bundy hoped EOD education and training would be altered to exploit these traits and to anticipate and accommodate the learning capabilities of typical EOD technicians. It is important for instructors to identify and understand these common strengths and

weakness among bomb disposal technicians. This information will allow them to make adjustments to the training environment to maximize potential in preparation for real-life EOD situations. This information can also be useful to the EOD community in more accurately selecting the proper candidate for EOD training.

The results on the LSI and MIDAS tests were very similar for bomb technicians. On the LSI test, bomb technicians showed strong commonalities in these three areas (Bundy & Sims, 2007).

1. Conditions for Learning: Over 75 percent of bomb technicians sampled showed high preference for:
  - *Detail:* Requiring specific information on assignments, requirements and rules
  - *Authority:* Desiring classroom discipline and maintenance of order
  - *Organizational:* Wanting course work to be logically and clearly organized with meaningful assignments and sequence of activities
  - *Competition:* Desiring comparison with others; needing to know how one is doing in relation to others
2. Expectation-for-Course-Grade: Roughly 72 percent reported a high Expectation-for-Course-Grade.
3. Learner Typology: Technicians had preferences common to the *Social/Applied* and *Independent/Applied* categories. The categories are similar in that all prefer opportunities to work in situations that approximate real-world environments while *Social* learners seek work with others and *Independents* prefer to work alone in a self-selected path toward a goal.

The MIDAS test (developed by Branton Shearer of M.I. Research and Consulting, Inc., in 1987) provided unusual results (Shearer, 2007). Typical groups are generally evenly distributed among the various types of intelligences (Bundy & Sims, 2007). However, bomb technicians consistently rated themselves significantly lower as compared with other populations. Of the eight types of intelligence strengths measured by MIDAS (Linguistic, Logical, Spatial, Kinesthetic, Musical, Interpersonal, Intrapersonal, and Naturalistic), most technicians rated themselves strong on:

1. Interpersonal Intelligence (44%): The potential for working with others, as used in understanding people, leading and organizing others, communicating and resolving conflicts.

2. Intrapersonal Intelligence (42%): The potential for understanding ourselves as used in recognizing one's own strengths and weaknesses and setting personal goals.

In contrast, EOD technicians scored themselves low on:

3. Musical Intelligence (24%): The potential for thinking in music; for hearing, recognizing and remembering patterns as used in singing, identifying sounds and remembering melodies and rhythms.

The results of the two tests, MIDAS and LSI, show that some characteristics among bomb technicians are correlated. The high score on the *Intrapersonal Intelligence* portion on the MIDAS exam and the high *Expectation-for-Course-Grade* on the LSI exam show that a person with high self-efficacy, or self-worth, would have high expectations of receiving a good grade. Exploiting this information in candidate selection and in training would logically increase the success of current and future bomb technicians (Bundy & Sims, 2007).

Both the military and non-military bomb technician communities seek to attract, train, and retain the most physically, mentally, and emotionally capable individuals who can perform while under the pressures of the bomb disposal job (Bundy & Sims, 2007). While both communities' recruitment efforts are generally successful, there are still manning shortfalls in part due to the attrition rate at EOD School. Using the common characteristics among bomb technicians identified by Bundy and Sims, these mismatches between learning style preference or intelligence strengths of an EOD candidate and those EOD technicians who have been successful in the field can be corrected to minimize attrition and manning shortfalls while still maintaining the quality of the student. Due to the dangers of the job, it could be disastrous to lower the standards to increase throughput.

Bundy and Sim's research is also useful in pre-selecting EOD candidates. Currently, the Navy uses the ASVAB, a physical fitness test, and an interview to select candidates into the EOD community. Other military services use less screening to select their candidates. In conjunction with these screening procedures, commonalities should be referenced before selecting a candidate. "If the student displays the common learning

style preferences and intelligence strengths shared by successful EOD technicians, he or she may be more likely to graduate than a student who does not possess the similar characteristics” (Turse & Ritland, 2009).

In their article, “Psychological and Physical Performance Factors Associated with Attrition in Explosive Ordnance Disposal Training,” authors Hogan, Hogan and Briggs wrote a study designed to predict performance in EOD training. Supported by the Naval Medical Research and Development Command, they conducted three studies following EOD students through the different phases of EOD training. As with Bundy and Sims, the authors wanted to develop measures of identifying the best candidates for the bomb disposal community and to reduce the high attrition rates due to misguided recruiting. The first study analyzed the psychological factors of the students at EOD School. The second study analyzed the physical performance factor as a predictor of success. Finally, the third study analyzed both the psychological and physical factors associated with the Navy’s twelve-week second-class diver course. The overall purpose was to create a set of selection procedures and guidelines for recruiting potential EOD candidates (Hogan, Hogan, & Briggs, 1984).

In the first study, non-cognitive tests were given to a sample of EOD students while undergoing training at EOD School. There were four of these tests:

1. CPI (California Psychological Inventory)—the most fully validated measure of normal personality
2. HPI (Hogan Personality Inventory)—assesses six factors associated with status and popularity in everyday life: Intelligence, Adjustment, Prudence, Ambition, Sociability and Likeability
3. SDS (Self Directed Search)—the standard vocational preference battery
4. ASVAB (Armed Service Vocational Aptitude Battery)—the primary cognitive battery used in the Armed Services

The results of the multiple tests revealed that EOD technicians were realistic, investigative, intellectual, self-assured and had social interests. These characteristics paralleled the profiles of athletes, engineers, pilots or technicians. People who deviate from this profile, such as artists or musicians, may not be successful in EOD training and

would be at a high risk for attrition. However, candidates possessing these traits will be more likely to successfully complete the rigors of EOD training (Hogan, Hogan, & Briggs, 1984).

The research team also determined that the use of vocational preference and non-cognitive measures are highly reliable predictors of academic success at EOD School. The team also found that the ASVAB was of little utility and highly inaccurate in selecting candidates who would ultimately graduate from EOD training (Hogan, Hogan, & Briggs, 1984).

In their second study, Hogan et al. analyzed the physical aspects of Navy EOD training. In 1982, pre-conditioning training and dive school accounted for 70% of the total attritions in the entire Navy EOD training pipeline (Hogan, Hogan, & Briggs, 1984). In 2008, pre-conditioning training accounted for 50% of attrition, dive training accounted for 20% of attrition and EOD School for 30% of Navy attrition (Getman, 2009). Hogan et al. identified seven dimensions that provide a complete coverage of physical strengths needed for job performance in any demanding field: Muscular Strength, Muscular Power, Muscular Endurance, Cardiovascular Endurance, Flexibility, Balance, and Neuromuscular Coordination.

The results of the physical study found that an extensive array of measures is necessary to predict performance in complex training programs. The researchers had to administer 26 different physical tests to lead to accurate prediction of successful dive training. They found that, of the seven dimensions, muscular strength was not a predictor of performance. Similarly, height, weight and body fat were not accurate predictors either. These findings suggest that successful performance in an arduous physical job is not related to physical size or strength. The best predictor, they determined, was cardiovascular endurance (Hogan, Hogan, & Briggs, 1984).

Their final study combined multiple elements, such as psychological, cognitive, physical and manual dexterity, to predict successful completion of dive training. They found that attrition in dive training is due to a specific set of factors. Students do not tend to fail due lack of cognitive competency; they did not typically attrite for academic



reasons. Instead, personal and physical factors were the primary reasons for training failures. The psychological tests determined that students who were successful in dive training were well-adjusted, self-confident and mature, as well as hard-working and achievement oriented. Those who were not successful were categorized as immature, anxious and self-doubting. The physical tests most predictive of successful performance were cardiovascular and muscular endurance. Therefore, candidates must be able to persist in physical activity while withstanding fatigue to graduate from dive training (Hogan, Hogan, & Briggs, 1984).

Overall, Hogan, Hogan and Briggs identified predictors for success during the academic portion of EOD School. They have also determined who will fail out of pre-conditioning training and who is at risk to attrite during dive training. By testing EOD candidates with the SDS and HPI, and incorporating cardiovascular endurance runs into the screening process, program managers can significantly reduce attrition in the EOD community (Hogan, Hogan, & Briggs, 1984).

The extensive research done in this field, dating from 1982, indicates that EOD attrition is high. However, the problem lies not in the demanding and arduous EOD training curriculum, but in the candidates selected for training. There exists a percentage of the population who are not cut out for the highly stressful, physically demanding and mentally challenging job of bomb disposal. Those people must be weeded out of the process. However, if recruiters can use the research findings that successful EOD technicians share similar traits, while tweaking the screening process to include specific psychological and physical factors to select the appropriate candidates from the beginning, standards in the community would not suffer, manning would increase and the military would experience cost savings.

The first task considered by Turse and Ritland (2009) was to compare the data from what the school provided us to what is shown in the school's instruction. We found our data to be slightly different from the data shown in the school's instruction, as referenced in differences in sample sizes. This is important to note, as our calculations and analysis in the previous study were based on the data the school provided. For this thesis, the same data provided by the school will be used to conduct the analysis.

## **A. NAVSCOLEOD INSTRUCTION 5420.1U**

Until 2004, a student who received an academic setback would sit through a full academic review board. However, as previously stated, the time and cost to conduct these boards grew as the military increased the number of students being sent to EOD School. This increase led to the development of the Training Officer (TO) Setback. This new policy has helped minimize time lost to instructors and students by conducting interviews and producing a student prediction for graduating based on historical data. This process would essentially replace an ARB and determine the outcome of a student, whether he or she should be set back or dropped from training upon the failure of his or her first retest. The model identifies how likely a student is to complete the entire training curriculum at the point of his or her initial setback. However, after the initial setback, if a student fails a test and a retest again, a full ARB must be convened.

The school's instruction, NAVSCOLEODINST 5420.1U, describes the Training Officer Setback policy:

The decision tool will improve training efficiency without compromising standards. The decision tool uses GPA and the first setback area to predict graduation. The tool predicts successful completion of training for 95% of graduates who experienced a setback, and the tool is far more accurate overall than the traditional ARB process.

The tool has an additional value in that changing the decision threshold allows it to predict nearly 100% of graduates while keeping false alarms in check. This decisional feature enables the tool to respond to forward demand signals more efficiently than the traditional ARB process.

Process. The student's first Academic Review Board/Training Officer Setback will consist of reporting to the Training Office in proper service dress uniform. The TO, Assistant Training Officer or Training Leading Chief Petty Officer will interview the student. The justification to remove from or continue with training may be based upon the student's grade point average and division recommendation where the student would be set back in. If the student is not at or below the minimum allowed GPA for the specific division the training office may grant a TO setback if warranted. The student will be recommended for drop if the GPA is less than that determined by the graduation prediction model (NAVSCOLEOD, 2008).

This process simplifies the decision to keep or remove a student from school. If the student has a graduation score above the historical average, the student has a higher probability of completing training and graduating from EOD School. Conversely, if a student's graduation score is below the cutoff, it is less likely that the student will succeed in school and therefore should be dropped from training. The instruction further describes the policy:

Statistical Model. The model will only be used for the area of the first setback to help determine if a student has the ability to complete NAVSCOLEOD objectives, and will not be used to address any further academic failures.

Annual Statistical Certification. The statistical data used in lieu of the first ARB will be checked on an annual basis using the first class convening in the new fiscal year. This class will be used to ensure the statistical method is still valid. Every student within the class will be given an ARB vise using the model for first setback situations. Each student that is given an ARB will be compared to the model. Using the table below, the model non-graduate predictions should not differ by more than the number in the right hand column.

The Training Officer will be responsible for maintaining statistical validation data for this model. Additionally, the Training Officer will coordinate periodic Technical Training Acceptance Board review of the ARB instruction and the annual statistical validation results to ensure the process is producing desired results. Subsequent failure of retests will result in an ARB (NAVSCOLEOD, 2008).

Table 1. Allowable Model Accuracy Tolerance Based on Sample Size

Class Sample Size (Number of Students)	Students who Graduate that the Model Predicted Would Not
12–20	4
21–24	5
35–44	6

While the model seemed like a partial solution to the school's problem, its usage underwent some scrutiny during the implementation phase. According to Chapter 3, Section 6 of NAVEDTRA 135B, students who are enrolled in Class "A" and "C" schools will only be academically dropped from training as a result of an ARB recommendation.

The Navy divides its training into categories to identify the certain types of training it provides. Class “A” schools provide basic technical knowledge and skills required for a rating and further specialized training. Class “C” schools provide advanced knowledge, skills and techniques to perform a particular job in a billet (NETC, 2009). Chapter 3, Section 6 of NAVEDTRA 135B also states that administrative procedures that result in “automatic” drops or setback are not authorized. At first glance, the model violates the Navy’s education regulations. Therefore, the Human Performance Center Detachment at the Center for Explosive Ordnance Disposal and Diving conducted an assessment to determine if this was in fact the case. After thorough review, they concluded the following:

1. NAVEDTRA 135 does not dictate how an ARB decision will be made.
2. The use of a decision tool does not preclude normal chain-of-command routing for CO approval.
3. The decision tool is an unbiased recommendation.
4. The decision tool is a better overall predictor of graduation outcomes than the traditional ARB process.
5. Decision tool output will be forwarded to the International Military Student Manager when an international military student is under review (Swiergosz, Aaberg, & West, 2005).

The Performance Center also stated:

NAVSCOLEOD has collected data over a two-year period from FY04-FY05 to develop the decision tool that predicts successful completion of training. These efforts produced the following regression equation:

$$\text{Graduation Score} = -4.585 + 0.057 \times \text{GPA} + 0.032 \times \text{Setback Area}$$

Where -4.585 is the y-intercept (the point at which the regression line crosses the y-axis), 0.057 is the coefficient for GPA and 0.032 is the coefficient for setback area (variable that represents the first test failure area). Expected outcome probabilities are show in Table 1 when the threshold for predicting graduation (no-yes; 0-1) is set at 0.5.

As shown in Table 2, the ARB decision tool is a robust predictor of graduation (95%); less than 5 % of students who actually graduate will be “missed.” It is also evident from Table 2 that the student receives a “benefit of the doubt” from the decision tool in that, successful completion of training is predicted 27% of the time when a student will actually fail (false-alarm).

In other words, after setback students have either graduated or failed, the school looks back and compares their outcome to their prediction score.

The decision threshold can be set to achieve different outcome probabilities. For example, setting the decision threshold at 0.4 is expected to predict nearly all occurrences of graduation and elevate the false-alarm rate from 27% to 66%.

Table 2. ARB Model Probability Matrix

Model Prediction	Graduation	
	Yes	No
Yes	.95	.27
No	.05	.73

\*FY04-05 data (n = 1166). Decision threshold = 0.5.

The outcome distributions in Table 3 represent data collected during FY05 validation. These outcomes parallel the expected probabilities in Table 2.

Table 3. FY05 ARB Model Validation

Model Prediction	Graduation	
	Yes	No
Yes	208 (97%)	15 (29%)
No	7 (3%)	36 (71%)

\*Standard decision threshold = 0.5. Data are from a different sample (n = 266) than data used to derive the model. Actual counts are shown with graduation outcome percentages in parentheses.

The outcome distributions in Table 4 represent data collected during FY05 validation when the decision threshold was set at 0.4. The “hit” rate is parallel to the expected probabilities (100%) and the false alarm rate was significantly lower (49%) than the expected (66%)  $\chi^2(1) = 6.38, p < .01$ .

Table 4. FY05 ARB Model Validation With an Adjusted Decision Threshold

Model Prediction	Graduation	
	Yes	No
Yes	215 (100%)	25 (49%)
No	0 (0%)	26 (51%)

\*Decision threshold = 0.4. Data is from the same sample (n = 266) used to validate the 0.5 decision threshold (Table 2). Actual counts are shown with percentages in parentheses.

The Performance Center expressed concerns about the model that:

1. Stakeholders must determine what constitutes a significant deviation from the expected model probabilities listed in Table 2.
2. The only decision thresholds that appear to be useful are 0.5 (default) and 0.4 as previously mentioned. Setting the decision tool at 0.4 is expected to yield higher “hit” (100%) and false alarm rates (50%). Force demand signals and the cost of false alarms will presumably dictate the decision threshold over a designated time period.

The decision tool was implemented for three main reasons—to reduce man-hours associated with the ARB process, avoid training costs associated with academic failures and enhance the ability to meet force demand requirements (Swiergosz, Aaberg, & West, 2005). In use for the past four years, the model has been successful in meeting these goals. In the six-month period from October 2008 to March 2009, 306 students had their first academic setback. Using the model, the Training Officer was able to “by-pass” the ARB 277 times, convening only 29 actual review boards. During these six months, this new procedure saved approximately 831 man-hours for instructors and students. The problem, however, is that the model is five years old (the statistical GPA the model is based on reflects old data from 2004 and 2005). Turse and Ritland’s (2009) research in *GPA as a Predictor of Success in EOD Training* shows that data over the past few years has shifted slightly out of the school’s instructional standard. NAVSCOLEDOINST 5420.1U claims that the model predicts successful completion of training for 95% of graduates who experienced a setback, and that the model is far more accurate overall than the traditional ARB process. Based on updated student data from 2004–2008, the model

predicted 94.1% would graduate and 5.9% would fail. (Data from 2009 were not used, as many of those students are still in school.) Although this is not within the specified requirements of NAVSCOLEODINST 5420.1U, the numbers are similar; stakeholders must determine what constitutes a significant deviation from the expected model probabilities listed in Table 2.

A problem with the data analysis in the school's instruction is the perspective of the information. As was discussed by Turse and Ritland (2009), the school takes a "backward-looking" approach to decide the accuracy of the prediction model. This approach forces the user to categorize everyone into two groups, graduated and did not graduate, and then tests the success rates in those outcome groups (Turse & Ritland, 2009). To illustrate, Table 5 was taken from the data gathered from the years 2004–2005. The model prediction is on the far left vertical column of the table; the graduation outcome is on the top row. The school's view of this data is presented in Table 6. The two groups are designated Graduated "Yes" (graduated) and Graduated "No" (did not graduate).

Table 5. 2004–2005 Model Prediction vs. Student Outcome (Threshold = .5)

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	366	132	498
No	40	89	129
Total	406	221	627

Using the two vertical columns of "Yes" and "No," if each of the values is divided by the totals in the bottom row and converted into percentages, the result is the backward-looking table (Table 6).

Table 6. 2004–2005 Backward-Looking Analysis

<u>Prediction</u>	<u>Graduated</u>	
	Yes	No
Yes	90.1%	59.7%
No	9.9%	40.3%
Total	100%	100%

Figure 1 further illustrates this “backward-looking” concept.

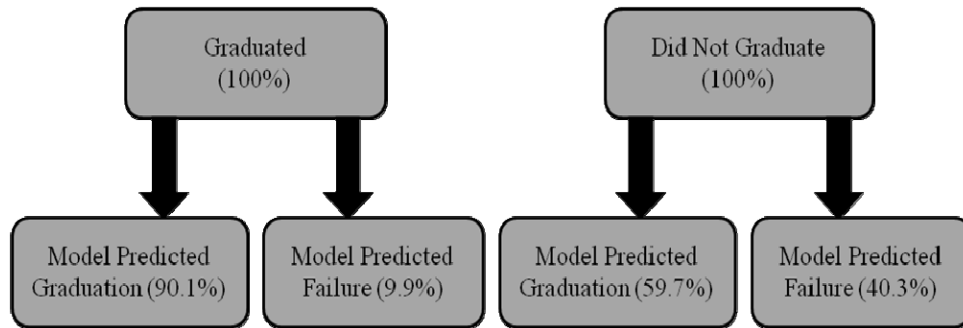


Figure 1. Backward-Looking Analysis

Once the groups are divided, the school’s instruction separately focuses on each group and determines at what the corresponding model scores were at the time of academic setback. In the graduation group, 90.1% of students had model scores greater than .5, so these students’ outcomes would have been correctly predicted at the time of the setback. Simply put, the model correctly identified 90.1% of the graduates. Among graduates, 9.9% had scores smaller than .5. Simply put, the model gave low scores to 9.9% of the graduates.

Among non-graduates, 59.7% had model scores above .5. Simply, the model kept 59.7% of students who would eventually fail out. Although this number may seem alarmingly high, it offers the student the “benefit of the doubt” by allowing him or her to continue training. Finally, among non-graduates, 40.3% had model scores below 0.5. The model correctly identified 40.3% of eventual non-graduates (Turse & Ritland, 2009).

As an alternative, Turse and Ritland (2009) have strongly recommended using the “forward-looking” approach. This approach allows the user to examine the prediction of the model without knowing the outcome in advance. As decision makers will not know a student’s outcome at the time of academic setback; it is more logical to use this perspective. In other words, once the prediction is made (graduate or will not graduate), the student goes into one of the two groups. The accuracy of the prediction is determined



when the outcome actually is known. To illustrate this, Table 7 was computed using the prediction “Yes” row and the prediction “No” rows and dividing the values by the row totals in Table 5.

Table 7. 2004–2005 Forward-Looking Analysis

<u>Prediction</u>	<u>Graduated</u>		<u>Total</u>
	Yes	No	
Yes	73.5%	26.5%	100%
No	31%	69%	100%

Figure 2 further illustrates the layout to this analysis.

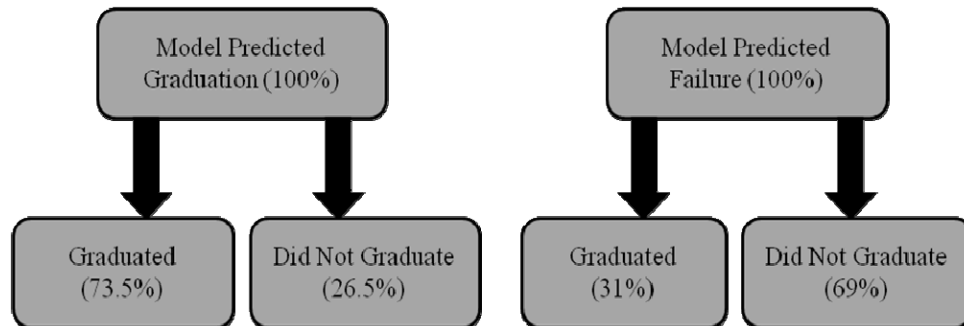


Figure 2. Forward-Looking Analysis

This approach is a more logical progression as the setback occurs first, followed by the outcome of graduated or not graduated. For example, in calendar years 2004 and 2005, 129 students were predicted to fail EOD school. However, 40 of those students (31%) went on to graduate. Conversely, 89 of those 129 students (69%) failed as predicted. Of the 498 students who the model predicted to graduate, 132 (26.5%) would have failed while 366 (73.5%) would have been predicted correctly. As the School’s instruction states, inaccuracy is more acceptable when the model predicts a graduation and the student fails (Turse & Ritland, 2009).

The problem with not understanding the two perspectives is a misunderstanding of the success of the model prediction. While the same initial data (from Table 5) was used to generate the percentages in the backward-looking and the forward-looking

analyses, the viewpoints are different (Turse & Ritland, 2009). EOD School's instruction only mentions one perspective and therefore misleads the model user. Turse and Ritland (2009) claim that the forward-looking approach is a more logical way to analyze the data.

The following are the differences that were shown in the MBA project. A difference in perspective means drastic differences in the percentages of success. The phrase "academic setbacks" was shortened to the term "setbacks."

1.     Correctly Predicting Graduation  
The Forward-looking Analysis correctly predicts graduation in 73.5% of setbacks who had a model score of .5 or above.  
The Backward-looking Analysis noted a model score of .5 or above in 90.1% of those who graduated.
2.     Incorrectly Predicating Graduation (False-Alarm)  
The Forward-looking Analysis falsely predicts graduation in 26.5% of setbacks who had a model score of .5 or above.  
The Backward-looking Analysis noted a model score of .5 or above in 59.7% of those who did not graduate.
3.     Correctly Predicting Failures (Non-graduates)  
The Forward-looking Approach correctly predicts failure in 69% of setbacks who had a model score of less than .5.  
The Backward-looking Approach noted a model score of less than .5 in 40.3% of those who did not graduate.
4.     Incorrectly Predicting Failures (Would-be Graduates)  
The Forward-looking Approach falsely predicts failure in 31% of those who had a model score less than .5.  
The Backward-looking Approach noted a model score of less than .5 in 9.9% of those who did graduate.

The last item is perhaps the most important difference between the two approaches. These two viewpoints present the different levels of error in the graduation prediction model. The forward-looking approach shows of all students the model predicts *will not* graduate, 31% of those students actually *will* graduate. On the other hand, the backward-looking approach shows of all students who graduated, 9.9% were predicted to fail. The model would have dropped these students from training, but given the opportunity they would go on to graduate and become EOD technicians. The backward-looking analysis claims to "miss" a much smaller number of students (9.9%)

then the forward-looking analysis (31%). It is important for NAVSCOLEOD to recognize this difference. By using the backward-looking approach, they believe their margin of error is small, while the forward-looking approach shows a much larger error rate (Turse & Ritland, 2009).

### **III. METHODOLOGY**

In their project, “Grade Point Average as a Predictor of Success in Explosive Ordnance Disposal Training,” Turse and Ritland (2009) asked the research question “Is the graduation prediction model still within the school’s standard?” They answered this question by analyzing the school’s database. This study asks the question “Is there a more accurate model by which to predict a student’s outcome?”

To address this question, we begin by analyzing the current model.

#### **A. MODEL REGRESSION**

The original model was generated using linear regression. This model was sufficient at the time it was developed. However, because the student’s outcome is binary (graduated/did not graduate), it would make sense to use logistic regression instead of linear regression.

#### **B. HISTORICAL DATA**

The incumbent model was developed by using the historical data prior to 2004; therefore the present thesis investigated the student data from 2004 to 2009. Although the School provided data from the past ten years (1999 to 2009), only the latest data has remained relevant. This is primarily due to EOD School steadily evolving in response to the changing global environment. In general, after a student graduates from EOD School, he or she joins an EOD team at one of the many EOD units worldwide and deploys to the Middle East roughly one year later. In order for these teams to maintain their high operation tempo, it is vital for EOD School to educate its students according to the latest threats and needs of the military. Therefore, older student data loses some of its validity to generate an improved model.

In their research using recent data, Turse and Ritland (2009) established that the model’s graduation prediction performance deviated slightly from the School’s standard of 95%. NAVSCOLEODINST 5420.1U says, “The tool predicts successful completion

of training for 95% of graduates who experienced a setback.” In other words, of all the students who graduated and had experienced a training setback, 95% had had a model score of .5 or higher at the time of the setback; therefore, the model predicted correctly that he or she would graduate (Turse & Ritland, 2009).

Their research also identified that the school’s instruction uses the wrong measurement of accuracy. As was discussed in Chapter II, this information is ambiguous. While 95% of the students who graduated were above the threshold of .5, this does not mean that the prediction at the time of setback (when the model users would make their decision) is 95% accurate. This measurement of accuracy does not suit the needs of the board nor the school. This is why the school should stop using this perspective. For this study, only the “forward-looking” perspective was used.

### **C. MODEL PREDICTORS**

Predictors were isolated to determine if any improvement could be made by adjusting their implementation. The model’s divisional input indicated a potential flaw. In the current model, each time a student completes a division, an equal increment of improvement is added to his or her overall chances of passing EOD School. This assumes each division is roughly equally difficult to pass. If this were true, it would result in roughly similar numbers of divisional setbacks from one division to the next. In their study, Turse and Ritland (2009) showed that this is not the case. Some divisions contained large numbers of first-time failures (e.g., air and ground division) while others had virtually none (e.g., WMD and Area VIII division). In response, the present study focused on the effects of changing the divisional input from numeric to categorical. This allowed each of the divisions to be isolated and each one’s contribution based on the student performance within that division to be weighed.

### **D. ADDITIONAL MODEL PREDICTORS**

The authors initially received a vast amount of historical data on each student at EOD School including age, branch of service, whether or not the student graduated, test scores, the division (if any) in which the student had a setback, and the reasons for

leaving school (graduation, academic drop, medical, etc.). Four years ago, the school decided that a student's GPA and setback division were sufficient, at the time of student setback, to accurately predict that student's future. This study identified whether other factors play a significant impact in improving the model.

### **1. Branch of Service**

In their study, Turse and Ritland (2009) analyzed the model results based on each military service. This information identified the differences among the branches of service in regard to academic performance. The authors showed that there is distinct separation in performance among the services, suggesting that branch of service, as a categorical predictor, might be related to the outcome in the model.

### **2. Interaction Among the Predictors**

The authors also considered whether the predictors showed interaction effects. This simply means that the effect of one predictor might be related to the value of another. For example, the effect of rank on graduation might be different in the Army than in the Navy.

## **E. COMPARISON OF MODELS**

The original model was then compared with the new models using two-way tables, ROC curves, graduation rate comparisons, and the Hosmer-Lemeshow test. In the next sections, these approaches are described.

### **1. ROC Curves**

ROC curves are graphs that compare sensitivity and specificity of a model with a binary outcome. (Fawcett, 2006) Sensitivity is defined as a "true positive rate," where specificity is defined as  $1 - \text{the "false positive rate."}$  A true positive rate (TPR) looks at how often the positive prediction was correct with respect to the set of positive outcomes and a false positive rate (FPR) looks at how often the positive prediction was wrong out of all the negative outcomes. Figure 3 is an illustration of a two-way table from which

TPR and FPR data are generated. In this, the positive values (graduated and predict graduation) are represented by the letter  $p$  or  $p'$ , and a negative values (predict failure and did not graduate) are represented by the letter  $n$  or  $n'$ . Capital letters denote the sum total of a given group and the prime represents the prediction. (Fawcett, 2006) The rows in the two-way graph (predicted outcome) are decided by some cutoff. When the observed value is greater than some  $p_0$ , then a positive prediction will result. For this given  $p_0$ , compute TPR and 1-FPR. This represents one point on the ROC Curve. The complete ROC Curve is traced out by varying  $p_0$ .

		Actual Value		
		$p$	$n$	Total
Prediction	$p'$	True Positive	False Positive	$P'$
Outcome	$n'$	False Negative	True Negative	$N'$
Total		$P$	$N$	

Figure 3. Prediction vs. Outcome Table

ROC curve analysis is a tool to compare multiple models using simple graphs. ROC analysis is a method to weigh the costs verses the benefits in the decision-making process. For example, the number of true positives may not be as important as the number of true negatives, or vice versa. Originally, the ROC curve was developed to locate enemy objects in the battlefield by electrical engineers and radar engineers during World War II. ROC analysis has also been used in medicine, radiology, and other areas like machine learning and data mining (Fawcett, 2006).

Drawing a ROC curve requires only the TPR and FPR data. Together, these two describe the results of the entire two-way table. As stated earlier, the TPR (sensitivity) determines the test performance for the predicted positive outcomes among all positive (graduated) outcomes. This is also known as the percentage of true positives ( $P(TP)$ ). Conversely, the FPR (1-specificity) defines how many predicted positive outcomes occurred among all negative (did not graduate) outcomes. This is also known as the

percentage of false positive (P(FP)). Figure 4 is an example of a ROC curve analysis showing the derivation of the ROC curve. Specificity and sensitivity are labeled on the x-y-axis respectively.

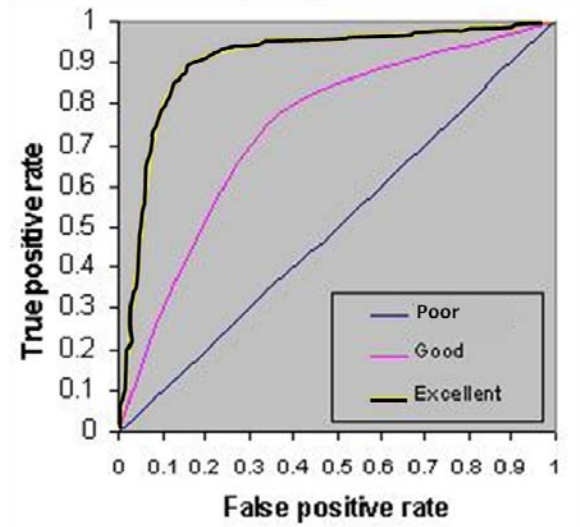


Figure 4. ROC Curve Analysis (From Lahanas, 2010)

A model with the best possible prediction would yield a point in the upper left corner of the ROC Curve graph. This represents 100% sensitivity (no false negatives) and 100% specificity (no false positives). This point is also called a *perfect classification*. In contrast, the model with the worst possible prediction (every guess being incorrect) would yield a point in the lower right corner of the ROC Curve graph. A point on the diagonal line, from the left bottom to the top right corners, represents a 50/50 random guess where 50% of the time the prediction is correct and 50% of the time the prediction is incorrect. This is the so-called *line of no discrimination*. Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold (Zweig & Campbell, 1993).

## 2. Hosmer\_Lemeshow Test

The Hosmer-Lemeshow test measures a lack of fit, usually for a logistic regression (Cook, 2005). This test involves partitioning predictions (in this case,



predicted probabilities of graduation) into (say) 10 equal-sized groups after sorting from least to greatest. After the groups are created, the observed number of successes (graduations) is compared with the expected number of successes (computed as the sum of predicted graduation probabilities). Table 8 represents an example of a Hosmer-Lemeshow table.

Table 8. Model Score Subsets

Grouping	$[x_0, x_1]$	$(x_1, x_2]$	...	$(x_9, x_{10}]$
Observed Successes	$O_1$	$O_2$	...	$O_{10}$
Expected Successes	$E_1$	$E_2$	...	$E_{10}$

We evaluate different prediction methods (models) by comparing each model's observed and expected successes. A better model would have predictions that more closely match the observed outcome.

Hosmer and Lemeshow proposed the following statistic by which to examine a model's goodness of fit. They suggest that if the resulting  $\chi^2_8$  statistic is associated with a p-value greater than .05, then the model is adequate. This approach cannot, however, be used on the original EOD School model since that model does not produce predicted probabilities of graduation. The following is the Chi-Squared test on Hosmer-Lemeshow (Cook, 2005).

$$G^2_{HL} = \sum_{j=1}^{10} \frac{(O_j - E_j)^2}{E_j(1 - E_j/n_j)}$$

$n_j$  = Number of observations in the  $j^{\text{th}}$  group


$O_j$  =  $\sum_i y_{ij}$  = Observed number of cases in the  $j^{\text{th}}$  group

$E_j$  =  $\sum_i \hat{p}_{ij}$  = Expected number of cases in the  $j^{\text{th}}$  group

### 3. Model Subset Comparison

As another way to evaluate models, rank-ordered model scores were compared in the following way. For each model, predictions were sorted and then arranged into the groups. We expect the smallest graduation rate in the tenth of the data with the smallest predictions. By computing the graduation rates within each tenth we can compare models when those models do not produce prediction probabilities. Table 9 illustrates the rank-ordered, graduation rate comparison concept.

Table 9. Rank-Ordered, Graduation Rate Comparison Example



	Current Model Performance	New Model Performance	Ideal Model Performance
Group 1	0.21	0.11	0
Group 2	0.32	0.21	0
Group 3	0.44	0.32	0
Group 4	0.57	0.64	1
Group 5	0.63	0.69	1
Group 6	0.67	0.74	1
Group 7	0.75	0.81	1
Group 8	0.81	0.85	1
Group 9	0.85	0.91	1
Group 10	0.91	0.96	1

THIS PAGE INTENTIONALLY LEFT BLANK

## IV. DATA ANALYSIS

The characteristics of the current model will be described in this chapter, followed by a discussion of this study's approach.

This study had roughly 1500 student observations. Our first focus was on the type of regression used by the model. The model currently uses a linear regression to generate a model score. These model scores can range from  $-1.33$  to  $1.33$  (Turse and Ritland, 2009). Generally, a linear regression is used to plot a relationship between predictor and a numeric outcome. An example of this is predicting the cost of a ship based on its weight. In this situation, EOD School has a binary outcome—every student either graduated or did not graduate. Since there is a binary outcome, the more logical choice would be a logistic regression. Logistic regression models the relationship between predictors and the binary outcome. In the current study, the authors were able to see the relationship between the predictors such as GPA, branch of service, and age and the graduation outcome. Our binomially distributed data is modeled by logistic regression in the following form:

$$\text{Graduation}_i \sim \text{Bernoulli}(p_i), \text{ for } i = 1, \dots, n$$

We modeled graduation in S-Plus® by using a generalized linear model (glm) of type binomial. This function provided us with the logits, or natural log of odds, of the unknown binomial probabilities. The function identified the coefficients that correlated with each predictor. These coefficients are additive effects on the logit associated with a one-unit change in each predictor (Hilbe, 2009). This following is an example of how the logits are generated.

Logistic Regression (logit)

$$\text{logit}_i = \beta_0 + \beta_1(\text{Age}_i) + \beta_2(\text{Branch}_i) + \beta_3(\text{GPA}_i) \dots$$

The logits can also be represented as the natural log of a probability ratio.

$$\text{logit}_i = \ln\left(\frac{p_i}{1-p_i}\right)$$

The following equation is the logit function of the logistic regression. Here, we can see how the probability is generated from the given logits. For simplicity, we used the S-Plus® function, `predict()`, to compute the probability associated with a logit.

$$\text{Logit Function}$$

$$p_i = \left( \frac{1}{1 + e^{-\text{logit}_i}} \right)$$

The probability derived from this equation gives the predicted probability of graduation. Figure 5 is a graph of the logistic function showing the relationship between logits and probabilities. The logits are represented on the x-axis and the probabilities are represented on the y-axis.

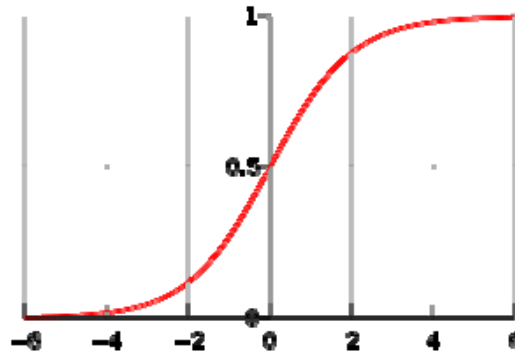


Figure 5. Logistic Function Graph

Next, we examined the predictors GPA and the Setback Division. The GPA predictor is reasonable, but there is a flaw in the implementation of the Setback Division. In the school's current prediction model,  $-4.585 + 0.057 \times GPA + 0.032 \times \text{Setback Division}$ , the student's first setback division (1-12) is multiplied by a coefficient (0.032) to generate the model score. Since the divisions increase in number from 1 to 12, a student is more likely to succeed, according to the model, after every division that student passes. The model score increases by an equal increment for each division. However, this equal increase assumes that each division is equally difficult.

We summed the total number of "first setbacks" in each division to examine the assumption of equal distribution across the years between 2004 and 2008. During these years, there were a total of 1485 students with setbacks. Figure 6 represents our results:

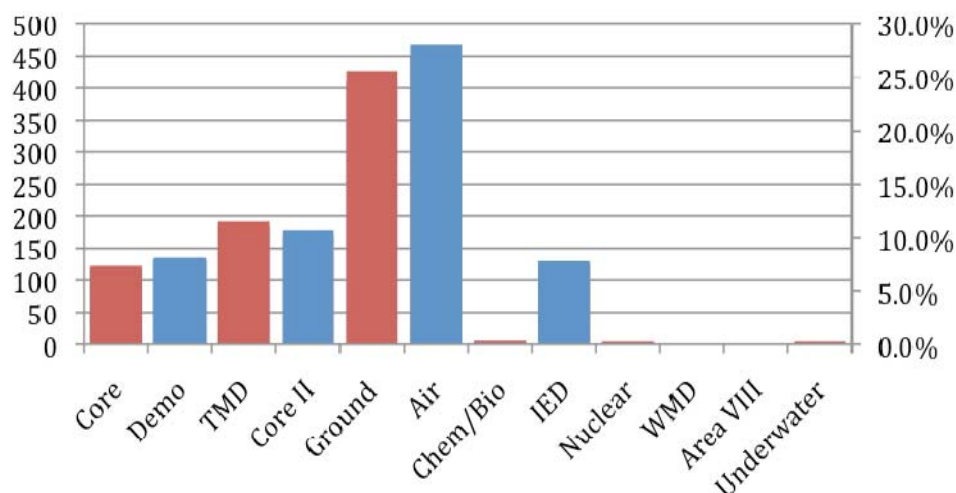


Figure 6. 2004–2008 Divisional Setbacks

There were setbacks in the Chem/Bio, Nuclear, WMD, Area VIII and Underwater divisions, although the percentages were so small that they are not clearly visible. As shown on the graph, the numbers of setbacks are not equally distributed. They generally increase through Air Division, and then drop off drastically. This data leads us to believe that the predictor, Division, should be categorical and not numeric. By changing the Divisional predictor to categorical, we can adjust the coefficient for each division, thereby accounting for the differences in the number of divisional setbacks. This change will, however, make the equation a little more complex than the current model.

THIS PAGE INTENTIONALLY LEFT BLANK

## V. RESULTS

In their study, Turse and Ritland (2009) concluded that the current model had, over the course of four years, slightly degraded from the School's prediction standard. They also determined that the School's instruction is misleading. The school's instruction claims the prediction model "predicts successful completion of training for 95% of graduates who experienced a setback." This is not to be confused with the prediction accuracy of the model at the point of setback. As was discussed earlier, the forward-looking analysis provides more useful information for decision makers.

Using only the forward-looking perspective, the results of this study show slight improvement in the overall probability of a correct prediction. The new model, however, is better for a number of reasons. The current graduation prediction model,  $-4.585 + 0.057 \times \text{GPA} + 0.032 \times \text{Setback Division}$ , is very simple. It requires only two pieces of information on the student to deliver his or her prediction. While the model is easy to use, it makes some basic assumptions. It assumes:

1. The linear regression is appropriate for this data;
2. The setback divisions contribute to the prediction of graduation in equal increments;
3. There are only two predictors that significantly impact the outcome;
4. There is no interaction between the predictors.

Each of these assumptions is in fact incorrect. As was stated previously, the type of regression used for the current model is generally used to plot a relationship between predictor and a numeric outcome. The data has a binary output—whether the student graduated or did not graduate. This suggests that the regression should be logistic. Therefore, a logistic regression to generate the results was implemented. With the knowledge that the present study's data would be better applied to a logistic regression, we addressed the rest of the assumptions before finding the final results of changing the type of regression.



Next, we address the setback divisions. As shown in Figure 6, the numbers of setbacks in each division differs significantly. This directed us to analyze the divisional component as categorical rather than numeric.

We then built a generalized linear model in S-Plus® and used a stepwise routine to determine which predictors should be in the model. The best set of predictors included Branch of Service, Setback Division, GPA and the interaction between these last two. Other traits including age, rank and gender did not play a significantly improve the model.

In analyzing the factors, we also looked for interaction among the predictors. While one factor may not play a significant impact on its own, when combined with another factor, it may improve the overall performance. We found that the interaction among the setback division and the student's GPA increased the prediction accuracy. The components of the improved model looked like the following:

$$\text{Graduation} \sim \text{Branch} + \text{Division} + \text{GPA} + \text{Division} \times \text{GPA} \text{ (interaction)}$$

Table 10 shows the resulting model. The divisions 7, 9, 10, 11, and 12 were removed because so few people failed in those divisions that we could not get an accurate coefficient to be used in the model. Instead, if a student failed in any one of those divisions, the coefficient for his or her setback division in the model equation is zero. This equation generates the logit, or natural log of odds. As discussed in the Chapter VI, to obtain the graduation probability, we use the logit function.

Table 10. New Model Equation

Branch+Div.+GPA+(GPA×Div.)
-10.851
0.000× USA
-0.145× USAF
1.278× USMC
0.187× USN
0.000× Division 1
-4.327× Division 2
-23.782× Division 3
-26.664× Division 4
-10.971× Division 5
-12.978× Division 6
-14.998× Division 8
0.120× GPA
0.049× GPA × Division 2
0.260× GPA × Division 3
0.297× GPA × Division 4
0.126× GPA × Division 5
0.158× GPA × Division 6
0.184× GPA × Division 8
Total = Graduation Score

For example, if an Air Force student had a setback in division 2 with a GPA of 90, that student's logit would be computed as:

Branch+Div.+GPA+(GPA×Div.)
-10.851
-0.145× USAF
-4.327× Division 2
0.120× GPA
0.049× GPA × Division 2
Total = Graduation Score

Using the example values, the logit comes out as:

Branch+Div.+GPA+(GPA×Div.)
-10.851
-0.145 × 1
-4.327 × 1
0.120 × 90
0.049 × 90 × 1
Total = -0.0753

This logit value is then entered in the logistic function equation to get the predicted probability:

$$p = \left( \frac{1}{1 + e^{0.0753}} \right) = .519$$

This probability equates to a 51.9% chance of passing EOD School.

If it is preferred that the branches of service be treated equally, the model in Table 11 is produced.

Table 11. New Model Equation Without Branch

Div.+GPA+(GPA×Div.)
-9.158
-4.859 × Division 2
-24.984 × Division 3
-26.941 × Division 4
-10.528 × Division 5
-13.071 × Division 6
-17.042 × Division 8
0.102 × GPA
0.055 × GPA × Division 2
0.275 × GPA × Division 3
0.301 × GPA × Division 4
0.122 × GPA × Division 5
0.160 × GPA × Division 6
0.207 × GPA × Division 8
Total = Graduation Score

We compared the current model to our proposed model (including the service variable) in four ways:

1. Two-way table comparison;
2. ROC curve comparison;
3. Hosmer-Lemeshow comparison;
4. Grouping graduation rate comparison.

The first comparison test was the two-way table comparison. The matrix is populated based on the prediction and the outcome. If a student was predicted to graduate and also graduated, he or she would be added to the top left box. The school's original model (with a .5 threshold) generated Table 12.

Table 12. 2004–2008 Current Model Prediction Outcomes

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	917	386	1303
No	58	124	182
Total	975	510	1485

As discussed, the backwards-looking perspective is not very useful to the model user, therefore we computed Table 13 using the forward-looking perspective.

Table 13. 2004–2008 Current Model Forward-Looking Analysis

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	70.4%	29.6%	100%
No	31.9%	68.1%	100%

This shows that when the model predicted a student would graduate, it was accurate 70.38 percent of the time. The table also shows that the model was accurate predicting a student would not graduate 68.13 percent of the time. The overall success

rate is computed by adding the correct guesses, 917 and 124, and dividing that by the total 1485 number of students. The current model has a 70.1 percent overall accuracy.

Using our new model,  $\text{Graduated} \sim \text{Branch} + \text{Division} + \text{GPA} + \text{Division} \times \text{GPA}$  (interaction), we generated a two-way table using the probabilities generated from the logistic function. We used a threshold of 0.5 although the two cutoffs are in different scales.

Table 14. 2004–2008 New Model Prediction Outcomes

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	847	284	1131
No	128	226	354
Total	975	510	1485

From this table, we computed the values in Table 14 using the “forward-looking” perspective.

Table 15. 2004–2008 New Model Forward-Looking Analysis

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	74.9%	25.1%	100%
No	36.2%	63.8%	100%

Comparing Table 13 (the current model) with Table 15 (the new model), we see an improvement in predicting graduation but a slight degradation in predicting failure. Overall, the new model was 72.3 percent accurate.

We also generated a two-way table without branch of service as a predictor (Table 16). Using the same format, we generated a two-way table using the probabilities generated from the logistic function. We kept the same threshold of .5.

Table 16. 2004–2008 New Model (Without Branch) Prediction Outcomes

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	852	294	1146
No	123	216	339
Total	975	510	1485

From this table, we computed the values in Table 17 using the “forward-looking” perspective. Overall, the accuracy of the model without service is 71.9 percent.

Table 17. 2004–2008 New Model (Without Branch)  
Forward-Looking Analysis

<u>Prediction</u>	<u>Graduated</u>		Total
	Yes	No	
Yes	74.4%	25.6%	100%
No	36.3%	63.7%	100%

Table 18 shows a table comparison between the accuracies using each model.

Table 18. Two-Way Table Overall Prediction Probability

Current Model	70.1%
New Model	72.3%
New Model (without Branch)	71.9%

Figure 7 shows Receiver Operating Characteristic (ROC) curves from the two models. Again, The blue line (the higher plot), generated by varying the threshold  $p_0$  is the ROC curve for the new model. The black line is the ROC curve for the school’s current model. While the improvement is not drastic, the new model outperforms the original model on almost every point of the curve.

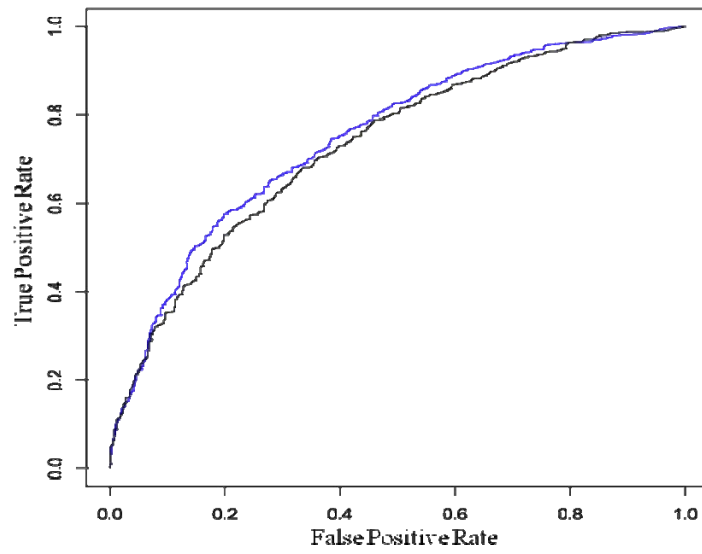


Figure 7. ROC Curve of New and Current Models

Figure 8 compares the new model without branch of service and the current model. Again, the new model is seen to afford at least a slight improvement.

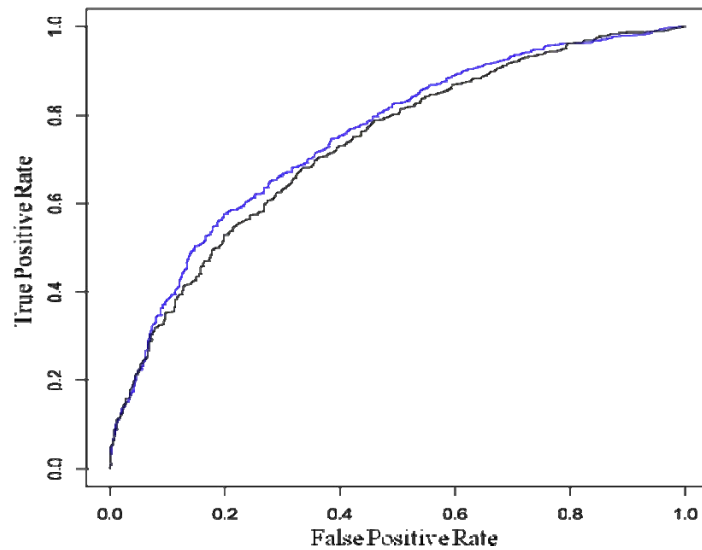


Figure 8. ROC Curve of New (Without Branch) and Current Models

We then we did a comparison using the Hosmer-Lemeshow test. As stated earlier, this approach cannot be used on the current EOD School model since that model does not produce predicted probabilities of graduation. We used it, however, to measure

the lack of fit for each new model. The two models we compared were a model with branch of service as a predictor and a model without the predictor. Figure 9 is the plot that was created using the new model with branch of service as a predictor. A good model has points that are close to the diagonal line.

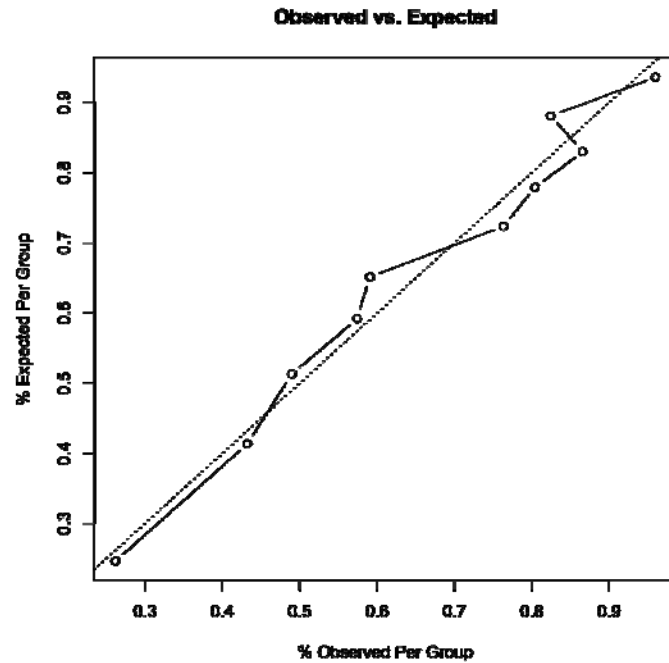


Figure 9. Hosmer-Lemeshow Plot of the New Model

We then ran the Hosmer-Lemeshow test on the model without branch of service as a predictor. The plot in Figure 10 was generating in the same way as Figure 9. This model plot performed better than the model with branch of service.



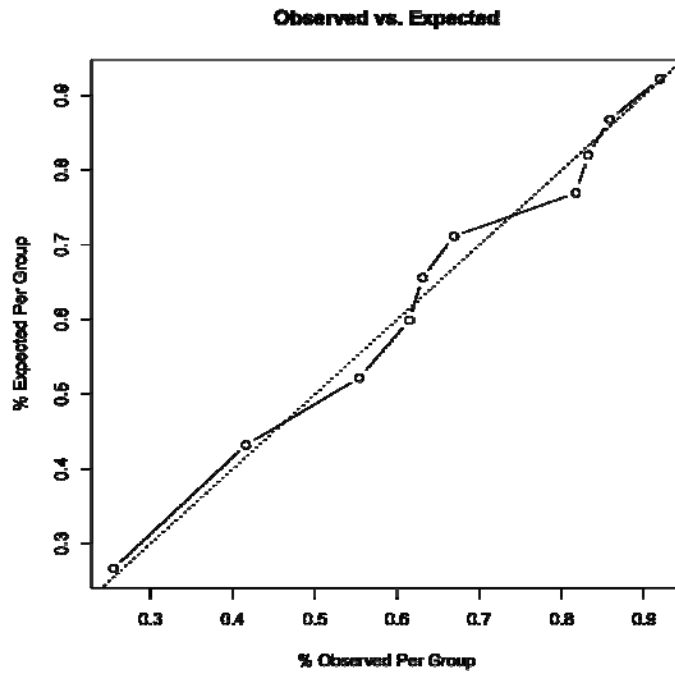


Figure 10. Hosmer-Lemeshow Plot of the New Model (Without Branch)

We then computed the Chi-Squared and p-value using the Hosmer-Lemeshow test. Table 19 shows the comparison between the two models.

Table 19. Hosmer-Lemeshow Test Comparison

	Chi-Squared	p-value
New Model	12.335	0.137
New Model (without Branch)	4.965	0.761

In the two-way comparison, the model with branch of service outperformed the model without it. Here, using the Hosmer-Lemeshow test, removing branch as a predictor actually improved the relationship between the prediction and the observation. However, in both cases the null hypothesis is believable.

Finally, we compared rank-ordered model scores. For each model, predictions were sorted and then arranged into ten groups. We expect the smallest graduation rate in the group with the smallest predictions. By computing the graduation rates within

groups, we were able to compare the current model, which does not produce prediction probabilities, to other models that do. The following table compares the graduation rates for prediction models.

Table 20. Rank-Ordered Graduation Rate Comparison

	Current Model Performance	New Model Performance	New Model (without Branch) Performance	Ideal Performance
Group 1	0.289	0.262	0.255	0
Group 2	0.48	0.432	0.415	0
Group 3	0.52	0.490	0.553	-
Group 4	0.591	0.574	0.615	1
Group 5	0.669	0.588	0.635	1
Group 6	0.685	0.765	0.664	1
Group 7	0.75	0.804	0.818	1
Group 8	0.812	0.866	0.832	1
Group 9	0.851	0.824	0.858	1
Group 10	0.919	0.960	0.919	1

In this comparison, we can see that the new model outperformed the original model slightly in most subsets.

Together with the other evidence of improvement this suggests that the new model is at least slightly better than the School's current model. Given the cost of model misidentifications, even a small improvement in accuracy can lead to a significant cost savings.

THIS PAGE INTENTIONALLY LEFT BLANK

## VI. CONCLUSION

The purpose of this thesis was to analyze the current graduation prediction model at NAVSCOLEOD and develop a better model using recent historical data. This thesis is a continuation to the MBA project conducted by Turse and Ritland (2009). Their project was designed to analyze the current model and verify that it has remained within the School's standard. NAVSCOLEODINST 5420.1U claims that the model predicts successful completion of training for 95% of graduates who experienced a setback. Using recent historical data, they found the model has slightly degraded in accuracy since its inception five years prior. Based on updated student data from 2004–2008, the model predicted 94.1% of the graduates would graduate.

The problem with School's instruction, however, is that it uses the wrong measurement of accuracy. To recapitulate, students at EOD School who fail an exam twice wait for an ARB to decide whether or not they should be retained. These boards are time-consuming and often ineffective. These boards are held for roughly 40% of the student population. As the global threats and situation have increased the demand for EOD technicians, the number of students attending EOD School has increased. As a result, the number of ARBs has increased as well, creating a bottleneck among students waiting for a board. As a solution to this problem, the prediction model was designed to aid the board members in the decision-making process. This is the prediction model's purpose. Yet, the school's instruction presents the model's accuracy based on an outcome-first perspective instead of prediction-first. This means that the instruction divides the groups based on outcome (graduated/did not graduate) and then determines the predictions of each group. Instead, the model's accuracy should be based on how accurate the prediction was. This requires the groups to be divided by prediction first and then the outcome in each group determined.

In this thesis, we set out to improve the current model. This process involved research into the school's historical data, investigation of the regression model, and computation of a new prediction model.

Since our data had a binary outcome (graduated/did not graduate), a logistic regression was deemed most appropriate for this data.

We examined factors outside the original model that play a significant impact in predicting a student's future at the time of academic setback. In our analysis, we found that the student's branch of service, GPA, and division all had a significant role on predicting a student's future. Other traits such as age and rank did not aid in accurate predictions.

We also changed the way one predictor was being implemented. The current model characterized the setback division as a numeric. Because of the different levels of difficulty among the divisions, changing setback division to a categorical variable allowed us to isolate each of the divisions and weigh each one's contribution based on the student performance within that division.

We also analyzed possible interaction between the predictors. Our results show that GPA and division have a positive interaction.

After running comparison tests on our models, our results show that our improved model, in which Graduation is modeled by Branch + Division + GPA + the Division×GPA interaction outperforms the other models on the two-way test. It also outperformed the current model on all other comparison tests. The new model without branch of service as a predictor performed the best on the Hosmer-Lemeshow test.

## **A. RECOMMENDATIONS**

Based on Turse and Ritland's (2009) MBA project and this thesis, we make the following recommendations to EOD School:

1. Stop using the "backwards-looking" perspective when measuring a model's accuracy;
2. Utilize the improved model developed in this thesis. Our results improved the model in four areas—using logistic regression, adding branch, categorizing division, and including one interaction. See Tables 9 and 10 for these models.

## LIST OF REFERENCES

- Andrea, T. (2009, March). *Assistant training officer naval school explosives ordnance disposal*. Eglin Air Force Base, FL: Department of the Navy.
- Atkinson, R. (2007, September 30). *The single most effective weapon against our deployed forces*. Retrieved January 11, 2010 from The Washington Post: <http://www.washingtonpost.com/wp-dyn/content/article/2007/09/29/AR2007092900750.html>
- Barnes, J., & Miller, G. (2009, February 18). *Obama orders more troops*. Retrieved January 11, 2010, from Los Angeles Times: <http://www.latimes.com/news/nationworld/world/la-fg-afghan-troops18-2009feb18,0,1590275.story>
- Booth, W., Colomb, G., & Williams, J. (2003). *The craft of research*. Chicago, IL: The University of Chicago Press.
- Brook, T. (2009, March 9). Coalition deaths from IED attacks soar in Afghanistan. Retrieved January 27, 2010, from *USA Today*: [http://www.usatoday.com/news/military/2009-03-08-IED\\_N.htm](http://www.usatoday.com/news/military/2009-03-08-IED_N.htm)
- Bundy, E., & Sims, R. (2007). *Commonalities in an uncommon profession: Bomb disposal*. Retrieved February 7, 2010, from Proceedings Ascilite Singapore: <http://www.ascilite.org.au/conferences/singapore07/procs/bundy.pdf>
- Cook, N.I. (2005, April 15). *Biostatistics*. Class Notes. Retrieved January 27, 2010, from University of Wisconsin: <http://www.biostat.wisc.edu/~cook/642.tex/notes0412.pdf>
- Fawcett, T. (2006). *An introduction to ROC analysis*. New York, NY: Elsevier Science Inc.
- Getman, R. (2009). *High attrition in EOD and ND pipelines*. Power Point Presentation. Panama City, FL: Naval Center for EOD and Diving.
- Hagey, K. (2007, August 21). *The bad guys are getting smarter*. Retrieved January 27, 2010, from CBS News: <http://www.cbsnews.com/stories/2007/08/21/iraq/main3188893.shtml>
- Hilbe, J. (2009). *Logistic regression models*. Boca Raton, FL: CRC Press.

- Hogan, J., Hogan, R., & Briggs, S. (1984). *Psychological and physical performance factors associated with attrition in explosive ordnance disposal training*. Bethesda, MD: Office of Naval Research.
- Center for American Progress (2007, March 19). Retrieved May 24, 2009, from Center for American Progress: [http://www.americanprogress.org/issues/2007/03/iraq\\_by\\_the\\_numbers.html](http://www.americanprogress.org/issues/2007/03/iraq_by_the_numbers.html)
- JIEDDO. (2008). *Joint improvised explosive device defeat organization FY 2008 report*. Washington, DC: Department of Defense.
- Military Factory. (2010). *2010 Military Pay Scale Chart*. Retrieved February 6, 2010, from Military Factory: [http://www.militaryfactory.com/military\\_pay\\_scale.asp](http://www.militaryfactory.com/military_pay_scale.asp).
- NAVSCOLEOD. (2008). *NAVSCOLEODINST 5420.1U Academic review board and student setback policy*. Eglin Airforce Base, FL: Department of the Navy.
- Navy EOD. (2009). Retrieved May 24, 2009, from Naval Explosive Ordnance Disposal: [www.eod.navy.mil/training.htm](http://www.eod.navy.mil/training.htm)
- Lahanas, M. (2010). *The receiver operating characteristic (ROC) curve*. Retrieved March 8, 2010, from Multiobjective Optimization with Evolutionary Algorithms: <http://www.mlahanas.de/MOEA/Med/ROC21.htm>
- Schmitt, E. (2009, November 30). *Obama issues order for more troops in Afghanistan*. Retrieved January 11, 2010, from The New York Times: <http://www.nytimes.com/2009/12/01/world/asia/01orders.html>
- Swiergosz, M., Aaberg, W., & West, H. (2005). *EOD academic review board preliminary fact finding*. Panama City, FL: CENEODDIVE.
- NETC. (2009, October 5). *NAVEDTRA 135B Navy School management manual*. Retrieved October 18, 2009, from Credential Opportunities On-Line: <https://www.cool.navy.mil/navy-edu-train-sys.htm#a>
- Turse, S. & Ritland, T. (2009, December 15). *Grade point average as a predictor of success in explosive ordnance disposal training*. MBA project, Naval Postgraduate School.
- Wehmeyber, M. (2007). *Explosive Ordnance Disposal. Power Point Presentation*. Arlington, VA: Headquarters U.S. Air Force.

Zweig, M. & Campbell, G. (1993). *Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine*. Retrieved March 9, 2010, from Clinical Chemistry: <http://www.clinchem.org/cgi/reprint/39/4/561>



THIS PAGE INTENTIONALLY LEFT BLANK

## INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center  
Ft. Belvoir, Virginia
2. Dudley Knox Library  
Naval Postgraduate School  
Monterey, California
3. Samuel E. Buttrey  
Naval Postgraduate School  
Monterey, California
4. Robert A. Koyak  
Naval Postgraduate School  
Monterey, California
5. Commanding Officer  
Naval School Explosive Ordnance Disposal  
Eglin Air Force Base  
Fort Walton Beach, Florida
6. Commanding Officer  
Center for Explosive Ordnance Disposal and Diving  
Panama City, Florida